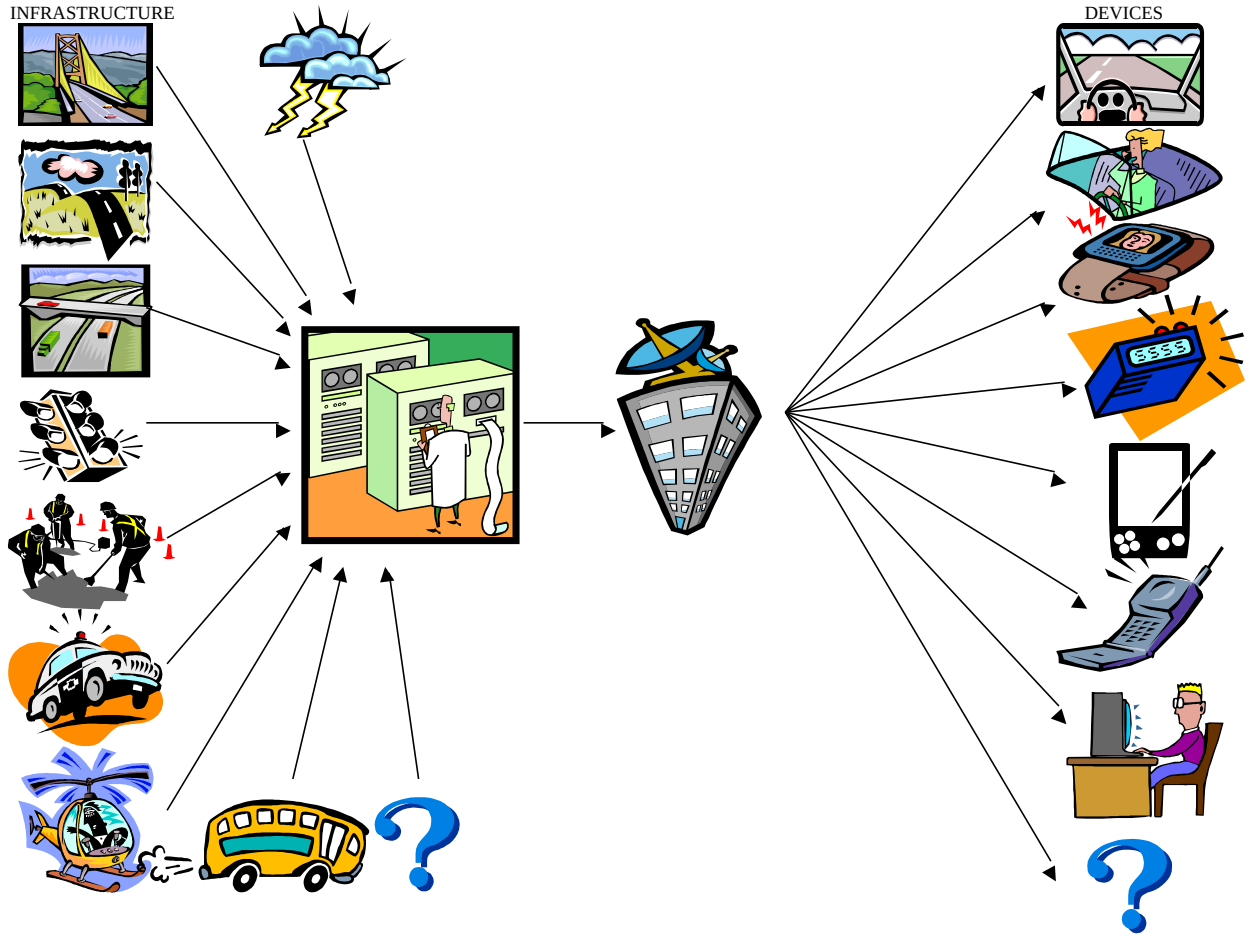
Closing the Data Gap: Guidelines for Quality Advanced Traveler Information System (ATIS) Data



Version 1.0
September 2000

ITS America

U.S. Department of Transportation

EXECUTIVE SUMMARY

In April 1999, ITS America's Advanced Traveler Information Systems (ATIS) Committee formed a Steering Committee to develop a national workshop on the subject of ATIS-related data collection and data quality. The goal was the creation of guidelines that can assist public agencies and private firms in generating and using data to support the expansion of ATIS products and services. Early on, ITS America's Benefits, Evaluation and Costs Committee joined in the effort by providing its official support and several additional committee volunteers. This enabled the collective effort to link discussions of data content and quality with the most recent evaluation findings of consumer preferences for the products and services enabled by such data.

Market research has been conducted in recent years to determine as best as possible consumer desires for ATIS products and services. In support of developing these guidelines, this research was synthesized and added to in key areas to provide the Steering Committee and workshop participants a complete picture of market needs and current deployment. In total, the research confirms that consumers' needs are tied to more and better data and that current data collection falls short of meeting the consumers' needs, in content, coverage and quality.

The Steering Committee has identified five reasons for publishing this document:

1. Raise awareness of the need for data collection planning
2. Increase amount of traffic data being collected
3. Increase quality of traffic data being collected
4. Increase recognition of the value of data
5. Encourage similar efforts for traffic management, transit management, and transit-related and rural traveler information data collection

The focus of these guidelines is limited to real-time or dynamic traffic-related information necessary to offer traveler information services envisioned in the near-term (roughly 5 years). At present, it seems that these are the data needs that must be addressed to accelerate market development. Limited access roadways and principal arterials are the subject of these guidelines.

The primary audience of this effort includes implementers and planners of traveler information services in both the private and public sectors. This audience will find the guidelines useful as a development, planning, and evaluation tool. The guidelines were formulated at a high level understanding, so executives and policy makers in the public and private sector will also find the guidelines useful as a communications tool.

To establish these guidelines, two separate – but interrelated – issues require consideration. First are data content issues. Content is broadly used in this document to define the type, coverage and quality of the data collected. Second are data access issues. Access describes the availability of data, usually being collected for other purposes, to organizations for use in creating ATIS products and services.

Data Content

Four types of real-time traffic data have been identified for data guidelines.

- Traffic Sensor Data
- Incident/Event Reports
- Images
- Road/Environmental Sensor Station Data

For the Traffic Sensor Data and Incident/Event Reports, consensus has formed regarding the attributes used to define the data type as well as the desired quality levels. The guidelines offer a baseline quality level, "good," and enhanced quality levels "better" and "best". If a data collection system meets the "good" quality level for all attributes, then the system is capable of supporting the envisioned ATIS products and services. Exceeding quality levels beyond good improves the data available and should improve the quality of the services that can be offered in the region.

The recommended minimum, or "good" quality levels are as follows (for description of the terminology used, please review the *Data Types, Data Attributes* and the *Data Quality Levels* sections of this document).

Data Type: Traffic Sensor Data

Attributes and Quality Levels:

- Nature: Limited Access Highways – Aggregated Point Data
  Principal Arterials – Aggregated Section Data
- Accuracy: < 15% error
- Confidence: Qualitative measure of suspicious data communicated along with the data
- Delay: < 5 minutes
- Availability: > 95% availability
- Breadth of Coverage: Limited Access Highways – Major Roadways
  Principal Arterials – Major Roadways
- Depth of Coverage: Limited Access Highways – Between Major Interchanges
  Principal Arterials – Between Major Arterials/Limited Access Highways

Data Type: Incident/Event Reports

Attributes and Quality Levels:

- Nature: Crashes, Breakdowns, or other unplanned vehicle stoppages; Planned or emergency roadway construction or maintenance; Disasters
- Detail: Reason, Location, Severity, Time
- Timeliness: < 5 minutes (for detection and verification stages, < 10 minutes total)
- Accuracy: < 15% error
- Confidence: Verified non-visual (the operator entering the information can not visually confirm)
- Breadth of Coverage: Limited Access Highways – All Roadways
  Principal Arterials – Major Roadways

Descriptions of the Images and Road/Environmental Sensor Station Data types contain possible attributes, but no attempt is made at present to define quality levels for these data types. In the future, the guidelines could be updated to include recommended attributes and quality levels as industry consensus emerges.

Data Type: Images

Possible Attributes:

- Breadth of Coverage
- Depth of Coverage
- Resolution
- Refresh Rate

Data Type: Road/Environmental Sensor Station Data

Possible Attributes:

- Nature
- Breadth of Coverage
- Depth of Coverage

Data Sharing

While who has access to data and under what terms is an extraordinarily complex issue for data collectors, the recommended quality level is straightforward:

- Access: Licensed or open access, via a documented interface

In summary, good data content and sufficient data access are required to support the development of ATIS products and services. Good data with limited or no access or inadequate data with sufficient access both result in less than optimal data being available to support ATIS products and services.

Relationship to Other Intelligent Transportation Systems (ITS) Areas

As a key element of an integrated ITS, ATIS data collection efforts must be coordinated with the existing data collection of other areas and systems, such as Advanced Traffic Management Systems (ATMS) or Advanced Public Transportation Systems (APTS). It is important for those developing a strategy for data collection to consider the needs of the various systems at work in their respective areas. This will help to foster a more coordinated regional effort toward data collection as well as a more cohesive plan for use and dissemination of the data from the various systems.

Applying the Guidelines

To assist organizations in implementing these ATIS data collection guidelines, an eight-step process for their application is outlined below. These steps are intended to serve as general "guidelines," and should be tailored to suit regional ATIS data-specific needs and issues.

1. Determine/define the region (boundaries) and stakeholders.
2. Outline "vision" for ATIS data to determine data needs and data priorities.
3. Determine what types of data are currently being collected from other ITS.
4. Select desired corridors for coverage.
5. Apply Minimum ATIS Data Collection Guidelines to selected corridors.
6. Consider local factors and traffic management needs.
7. Consider improved data quality goals.
8. Estimate cost and synchronize data collection strategy with funding profile.

Further Action

ATIS data collection guidelines are evolving. At present, this document contains sufficient detail only in a few areas where guidelines are actually desired. To reflect this, this document has been characterized as Version 1.0.

The Steering Committee recommends continued refinement and expansion of these guidelines to more broadly address all real-time data necessary to support qulaity ATIS products and services. Near-term focus should be placed upon images, road/environmental sensor station data, and real-time transit information.

The Steering Committee also recommends that these guidelines be reviewed, at minimum, by professionals from the traffic management, transit management, rural applications and archived data and transportation statistics fields to determine the possibility of expanding the guidelines to address real-time data collection needs of all ITS applications.

TABLE OF CONTENTS

INTRODUCTION

Advanced Traveler Information Systems (ATIS) has proven to be perhaps the most complicated area of ITS to deploy primarily because of the growing complexity of its "commercial architecture," the web of business relationships between companies, agencies, and individuals necessary to make something happen. In the early development of ITS in the United States, the vision for ATIS was relatively simple. Public agencies would collect data, and that data would be disseminated to devices in vehicles, homes, offices, wireless devices, and other media. In hindsight, it is clear that the difficulty of collecting good complete and timely data, transforming data into information, then packaging, marketing and communicating that information to people and devices was underestimated.

It is now evident that the traveler information market is not a single vertical market. Traveler information is evolving as another source of "content" for multipurpose devices in cars, homes, offices, and pockets – anywhere travelers and businesses seek and find information. In addition to providing some direct to consumer applications, traveler information service providers (ISPs) feed other vertical markets – such as cable TV, web portals, and paging services. Some services are subscription-based, others are advertising-driven. Still others are subsidized by the public sector.

With both the telematics and wireless data communications markets set to explode (and converge), the goal of widespread usage of real-time traffic and transportation information to improve the efficiency and quality of travel is poised to become reality. However, a fundamental issue has emerged in recent years: the sparseness of raw data necessary to support these services. How to increase the collection of data users want has been the focus of a year-long effort under the auspices of ITS America with active support from the U.S. Department of Transportation.

Focusing on data collection is a natural evolution of the initiative recently sponsored by ITS America and the U.S. Department of Transportation to assist regions to develop business models/plans for ATIS. As regions/agencies develop and determine their models/plans, they often seek input from ISPs and others, who would like access to data collected. The agency asks the ISPs: "what data do you want from us?" The response back from the ISPs is often: "whatever you have." Agencies do not really find this answer helpful in planning data collection strategies and priorities. To be fair to ISPs, the market is evolving so rapidly that they are unsure of the types of services that will become accepted, and of particular importance, the market model that will result. As a result, ISPs cannot tell an agency unequivocally if, for example, incident information on a major arterial is more or less important than travel times between interchanges on a limited access highway. Nor, for example, can they tell an agency if it is more important to get data that is less than 60 seconds old or if greater delay is acceptable if it enables funding to be used to increase data collection coverage.

Guidelines Development Process

In April 1999, ITS America's ATIS Committee formed a Steering Committee to develop a national workshop on the subject of ATIS-related data collection and data quality. The goal was the creation of guidelines that can assist public agencies and private firms in generating and using data to support the expansion of ATIS products and services. Early on, ITS America's Benefits, Evaluation and Costs Committee joined in the effort by providing its official support and several additional committee volunteers. This enabled the collective effort to link discussions of data content and quality with the most recent evaluation findings of consumer preferences for the products and services enabled by such data.

Through monthly teleconferences, the Steering Committee (see Appendix A for the list of members) developed the format and guided development of materials for the highly successful Advanced Traveler Information Systems Data Collection Guidelines Workshop, conducted February 9-10, 2000 in Scottsdale, Arizona. Output from the workshop is largely responsible for the content of this document. Key findings of the workshop were:

- ***Guidelines for quality ATIS data are desirable*** – Guidelines will aid public agencies (and emerging private data collection providers) in data coverage and investment planning. It will also enable data collectors to precisely communicate data collection capabilities to users of the data such as ISPs ("truth-in-labeling"). These guidelines must be designed as a guide, not strict rules.

- ***Need further refinement in classifying types of data, quality attributes for each type of data, and quality levels for each attribute*** – A group of workshop participants volunteered to assist in addressing these detailed issues. (Note: These volunteers, called the Data Guidelines Oversight Group, were instrumental in converging on the data types, attributes and quality levels. These volunteers are listed in Appendix A.)

- ***Guidelines for quality data go beyond ATIS*** – Over time, guidelines for ATIS data should be coordinated with traffic and transit management data collection requirements and other appropriate application areas to develop a complete guide to data collection to support transportation management and operations.

Since the workshop, efforts have been focused on drafting this Guidelines for Quality ATIS Data document.

Purpose of Guidelines

The Steering Committee has identified five reasons for publishing this document:

1. ***Raise awareness of the need for data collection planning*** – In conjunction with developing an architecture to guide an agency or region's ITS deployment, it is important to carefully consider the types of data to be collected and the performance targets to be set to supply the data necessary for the desired ITS applications.

2. ***Increase amount of traffic data being collected*** – Consensus has emerged that more data on roadway conditions is needed. Even in regions where data are collected, it usually is not collected on all roads of interest from an ATIS perspective.

3. ***Increase quality of traffic data being collected*** – Often data currently collected does not meet the quality levels identified in this document. This document aims to provide targets for improving in general the overall quality of data collected to support ATIS products and services.

4. ***Increase recognition of the value of data sharing*** – Anecdotal evidence suggests, and quantitative data seem to confirm, that a significant amount of data collected for other purposes is not available for use in traveler information applications. Raising awareness of the need for such data to support ATIS products and services could increase the sharing of such data.

5. ***Encourage similar efforts for traffic management, transit management, and transit-related and rural traveler information data collection*** –ATIS is often not the principal reason for data collection. Data quality is important for all ITS applications and, though this document offers guidelines focused on ATIS, it is recognized that more meaningful guidelines will be established when all ITS application constituencies develop guidelines such as these tailored to their application areas.

Guidelines Scope

The focus of these guidelines is limited to real-time or dynamic traffic-related information necessary to offer traveler information services envisioned in the near-term (roughly 5 years). At present, it seems that these are the data needs that must be addressed to accelerate market development. Limited access roadways and principal arterials are the subject of these guidelines.

This paper does not address static information such as road maps and transit fares/schedules/stops. These areas are important and are the subject of significant activity. However, the focus of this paper is concentrating on real-time or dynamic information. Road maps and transit fare/schedule/stop information changes, but not in a timeframe that can be considered close to real-time.

These guidelines do not address real-time transit information. Similar data guidelines for transit-related information should be developed as soon as possible. Consensus has yet to emerge as to what type of real-time transit-related information should be communicated to patrons. The Steering Committee encourages the transit industry, and potential consumers of transit data, to develop a consensus of what real-time information is important to travelers and how it might be used. When such a consensus emerges, data collection guidelines for transit applications should be developed and integrated with these highway-related data collection guidelines.

BACKGROUND

As mentioned earlier, ITS America's ATIS Committee has spent considerable energy focused on the topic of business models and business plans for ATIS. A document published by ITS America on the subject in 1998, *Choosing the Route to Traveler Information Systems Deployment*, emphasized that multiple models for deploying and operating traveler information systems are both possible and likely, and that the public agencies in a given region have the discretion to select the model that is most appropriate.

A common model emerging has at least some level of private sector involvement in providing ATIS products and services, often with some or most the data collection being performed by public agencies. Those who are using this model were the first to identify the possibility that an insufficient amount of quality data are being collected to support establishment of ATIS products and services in demand. The Steering Committee believes the issue of the possible lack of quality data is independent of the business model used. Thus, this document does not focus on specific models, nor does it endorse one model over others. For more information on business models, see the sources identified as references.

Regardless of business model approach, an updated vision of ATIS is emerging. This vision is more complex than the comprehensive traveler information solutions originally suggested around 1990. This vision is also more powerful. The vision contains three general, inter-related steps. Experience is telling us that these three steps require vastly different skill sets to accomplish:

1. Data Collection – This step involves gathering data that describes events and status of travel conditions

   Gathering useful, accurate, reliable and timely data to support ATIS applications is non-trivial. To date, data are usually collected for multiple purposes. ATIS is often a secondary or tertiary purpose subordinated to the principal of purpose of traffic or transit management, for example. However, significant technological and process innovation appears ready to make ATIS-centric data collection feasible in the near future.

2. Data Fusion – This step involves assembling data from multiple data collection sources into a useable, integrated form

   Sifting through multiple types of information requires skills to identify collaborative data and single out conflicting data from different sources so a unified picture of conditions can be available. Marketing the fused data stream to business partners is also a key part of this role. Data fusion is now occurring beyond regional levels. Some firms are providing a single data stream covering dozens of metropolitan areas, with goal of creating a national data stream for use by businesses seeking a national footprint.

3. Data Dissemination – This step involves communicating fused data to a consumer as a service, when, where and how a customer wants the information.

   Skills required to disseminate data include consumer/end-user marketing, account and customer-relationship management, and product planning. Travel information is likely to be blended with other useful real-time, perishable information, such as weather, stock quotes, sports, and news to provide maximum value to the user. Traveler information is emerging as a vertical content market in the emerging internet/wireless market space.

Market Research

Guidelines for ATIS data must be developed with the ultimate use of such data in mind. Also, the qualitative perception of a shortage of quality ATIS data required quantitative research to confirm or refute the hypothesis that a problem exists. Thus, the Steering Committee commissioned, and U.S. DOT funded, the development of three papers that collectively shed light on these issues:

- ATIS: Who are ATIS Customers?
- ATIS Data Collection Guidelines Workshop: What do ATIS Customers Want?
- ATIS: Private Sector Perceptions and Public Sector Activities

Like all market-related research the findings of these papers summarized here reflect the information available at the time of study completion. While the Steering Committee acknowledges that the research is based largely on the needs of metropolitan areas, and less upon the needs of inter-city, tourist, rural and freight travel, these findings seem to address the needs of these other market segments, based on anecdotal evidence.

Complete versions of these papers are available at ITS America's web site at http://www.itsa.org/ITSNEWS.NSF/4e0650bef6193b3e852562350056a3a7/eb34d0c5ae8ed31f8525690e00547667?OpenDocument.

Consumers and Demand for ATIS

Research of current ATIS users, as well as the general public that are not yet ATIS users has identified four factors that influence ATIS customer demand:

1. The regional traffic context: This includes attributes of the region, such as highway-roadway network and capacity, levels of traffic congestion, and future highway-roadway expansion plans. Prime ATIS markets appear to be highly congested regions that have limited build-out options, constrained alternate route possibilities, and frequent unpredictable traffic events (e.g., weather, crashes).

2. The quality of the ATIS services: This is at least as important as the level of network congestion. Information quality determines whether, how frequently, and with what level of confidence the traveler consults traveler information. Quality determines whether the information will meet customer needs with respect to personal benefit and value.

3. The individual trip characteristics: The trip purpose, the time of the trip in relation to peak congestion periods, trip length, and the particular route or route choices available to the individual traveler all have a significant effect on whether the individual will consult traffic information. To a limited extent, the availability and convenience of alternative mode choices for a given trip affects use of ATIS. Travel time flexibility, or lack thereof, is another determinant in the choice to consult traffic information.

4. The characteristics of the traveler: The fourth factor includes values and attitude characteristics of the ATIS user, or potential user. These characteristics are important determinants of user awareness, use patterns, behavioral responses, and valuation of ATIS.

Based upon numerous surveys, focus groups and research, eight consumer market segments have been identified that can be expected to cover approximately 90% of a region's population. The market segments help determine the potential market size for specific ATIS products and services, as well as the features of those services that are related to the data required to provide the service.

Of the eight segments, four seem to serve as the most likely candidates for early adopters of ATIS products and services:

1. Control Seekers: Like to plan ahead, desire to be accessible at all times, like using portable information devices, and want to predict travel time accurately

2. Web Heads: Most technologically savvy segment, high users of Internet, but low use of portable information devices

3. Low-tech Pre-trip information seekers: Prefer pre-trip information, and are less interested in new high technology gadgets.

4. Mellow Techies: Little interest in traffic conditions or trip planning, and little concern about being late, but high levels of computer and internet use.

The other four segments are:

5. Buyers of Value-added services: Low comfort with computers and Internet, may prefer customize information services.

6. Wired with Children: Younger, higher income, with more children in household, seeks convenience in information acquisition.

7. Trendy and Casual: Use pagers and cell phones, but express little interest in traffic information or time savings.

8. Male Techno-Phobes: Less comfortable with technology, less likely to change behavior, less interested in traffic information.

ATIS market segmentation based on attitudes and values related to control, time, travel, and technology successfully identifies much of the current ATIS customer market, differentiating ATIS customers from others with similar demographic characteristics.

Control-seekers dominate the ATIS customer market. These customers consult ATIS to save time, to use their time efficiently, to stay on schedule, and to stay informed. Control-seekers use information more intensively than the general population.

Technology has an important and complex role in ATIS. Web Heads comprise the second largest group of ATIS customers. However, their allegiance appears linked to the Internet media, and may or may not migrate to other information platforms as the web becomes more mobile.

Individuals in the low-tech pre-trip information seekers market segment had a low acceptance and comfort level with the Internet and web-based information. Consistent with this characterization, individuals belonging to this segment comprised much larger shares of the Transit Watch®, TrafficTV and TrafficCheck user populations. Nevertheless, this customer segment represents a large portion of the current ATIS customer pool, and can be expected to continue to demand good information services on low-tech media in the future.

Current Context

Drivers' point of reference for all traffic information is their personal experience with both local traffic conditions and radio traffic broadcasts. They generally rate their own experience as a reliable source of traffic information. But based on their experience with unreliable traffic information from the radio, some drivers do not believe that better, personally useful traffic information could exist. Other drivers believe that there's no alternative to traffic congestion and thus little value to ATIS. Therefore, new ATIS services are competing against drivers' personal knowledge of local traffic conditions, traffic broadcasts on the radio, and drivers' underlying belief that there's nothing ATIS could provide to relieve the situation.

Concurrently, consumers' expectations for advanced information services generally are very high. They've been conditioned by the Internet and a computing environment in which information services and electronic devices get faster, better, and cheaper very quickly. In the research and evaluation to date, we see a progression in the expectations and requirements of drivers as they become more experienced ATIS consumers.

Why do travelers use ATIS?

Washington State DOT traffic web site (believed to be the heaviest used real-time traffic web site) customers provide insight into motivations of use. These motivations are representative of most ATIS users in other regions. The answers to the questions are place in order of frequency:

Why use the web site?

- To assess traffic congestion on their route
- To judge the effects of incidents on their trip
- To decide among alternate routes
- To estimate their trip duration
- To time their trip departure

What action results from the information?

- Change route or time of departure maximizing for a faster trip time
- Change route or time of travel to reduce the stress of driving in congestion, perhaps lengthening trip distance or duration
- Adjust their expectations, listen to an audiotaped book, make phone calls, adjust appointments, and make alternative arrangements.

What benefits are perceived from use?

- Saved time
- Avoided congestion
- Reduced stress
- Avoided unsafe conditions

Critical features of a traffic-related ATIS service

The U.S. DOT ITS program fielded qualitative market research in 1996 on various traffic information concepts with drivers in congested regions. While the opinions of these drivers' were based on their experience of radio broadcast traffic information, their traffic information concerns have proven to be true of all drivers surveyed since.

- Accuracy of information
- Timeliness of information
- Reliability of information
- Cost to use
- Degree of decision guidance and personalization offered
- Convenience of access and speed
- Safety of operation

The content and quality of data collected will clearly affect the first five items of this list.

What do Information Service Providers Want?

ISPs as well as other private firms were surveyed in late-1999 to obtain their current assessment of the types of data needed to meet market demand and the content and quality of the data presently available. Again, while not every ISP or private company developing an ATIS product or service provided feedback, sufficient responses were obtained to give the Steering Committee confidence in the results.

Data Collection Priorities

The private sector is most interested in limited access highway data in the near term. The private sector ranking of particular types of information is as follows:

- Traffic speeds
- Incidents
- Road conditions
- Current and scheduled work zones
- Weather conditions

Data Quality Issues

The most frequently cited reason for insufficient data quality is inadequate geographic coverage. This arises mainly from incomplete data collection in metropolitan areas with multiple jurisdictions, particularly with respect to traffic speeds. Of the problems listed, geographic coverage and spatial resolution are the only ones for which it is possible to develop a national picture based on the public sector responses. To some extent the issues of update frequency and temporal coverage can be explored indirectly. The most common complaints, in order of frequency of citation are as follows:

- Inadequate geographic coverage
- Inaccurate information
- Insufficient update frequency
- Not timely enough
- Inadequate spatial resolution

What data do agencies presently provide?

The primary source for information on data collection and dissemination by public agencies is preliminary data from the 1999 metropolitan ITS deployment tracking database, funded by U.S. DOT.

Several conclusions can be drawn from the database that confirm the magnitude and nature of the gap in quality data collected for ATIS.

- Agencies that collect traffic sensor data on limited access highways (e.g. freeways, toll roads), collect such data on less than one-third of their centerline miles on average
- Most traffic sensor data are generated by "point" data sources, such as loop detectors, video image detectors and microwave radars
- Sensor data coverage on arterials capable of supporting ATIS applications is almost non-existent
- More incident data collection coverage exists than sensor data coverage
- For incident-related data collection, twice the number of agencies use closed circuit TV or automated algorithms that detect incidents from changes in traffic sensor data than use police patrols or reports from drivers using cellular phones
- Yet, agencies that use police patrols use them on over 70% of their mileage, agencies that seek help from motorists through cellular calls use them on roughly 80% of their mileage; in contrast, agencies using CCTV or automated algorithms do so on roughly 30% of their mileage
- In general larger regions have more agencies operating roadways
- The more agencies operating roadways in an area the more difficult to create the desired data in a region for ATIS use
- For a variety of reasons, much of the data presently collected by public agencies is not being provided to support creators of ATIS products and services.
- Public sector agencies do not believe providing traffic sensor data is as high of priority as compared to work zone, incidents, events and road and weather conditions.

These findings confirm, when compared to the guidelines to be presented in a later section of this document, that a gap in the type and quality of data necessary to support successful ATIS products and services does indeed exist and must be addressed to achieve the emerging vision of ATIS.

DATA QUALITY GUIDELINES

Definition

A prerequisite to the development of data quality guidelines is a clear definition and intent of what is meant by "data quality guidelines". Decomposing the phrase into its individual definitions provides a starting point. The following definitions were found at http://dictionary.msn.com:

> *Data* – information, for example numbers, text, images, and sounds, in a form that is suitable for storage in or processing by a computer
>
> *Quality* – the general standard or grade of something
>
> *Guidelines* – an official recommendation indicating how something should be done or what sort of action should be taken in a particular circumstance

For the purposes of this effort, "data" refers to real-time/dynamic road-related information to support traveler information services within the next five years. "Quality" has been generalized to paint a picture of the utility of data collected for traveler information services. The quality metrics, termed data attributes in this report, were designed with the intent of being technology independent so as to remain useful as the market matures. This report was developed via a consensus building process through workshops, meetings, teleconferences, and subcommittees representing a national audience of both public and private stakeholders. Consequently, these findings draw upon the experiences and needs expressed by those who collect, process, fuse, disseminate, market, and consume data in the United States. The guidelines are not prescriptive. Rather, the guidelines are a tool that encourages best practices as understood to date. The desire is that the guidelines will assist in accelerating the deployment of data collections systems necessary to support traveler information services.

Purpose

The Data Quality Guidelines are intended to be a useful tool for anyone involved in traveler information services. Implementers of traveler information systems will find the guidelines useful as a planning tool.

The guidelines also serve as an evaluation tool for those who have already implemented a traveler information system. Evaluators can use the guidelines as a national baseline to gauge the quality of the services provided.

Wherever possible, the guidelines have relied upon terminology and findings of existing standards efforts. As a result, the guidelines can be used as a communications tool between agencies and ISPs. The goal was to promote "truth-in-labeling" to foster more effective communications.

Since the guidelines are rooted in the needs of consumers and ISP preferences, implementers can use the guidelines to develop traveler information systems that are traceable to user needs.

Audience

The primary audience of this effort includes implementers and planners of traveler information services in both the private and public sectors. This audience will find the guidelines useful as a development, planning, and evaluation tool. The guidelines were formulated at a high level understanding, so executives and policy makers in the public and private sector will also find the guidelines useful as a communications tool.

DATA TYPES, DATA ATTRIBUTES

Two concepts are key to defining quality data: Data Types and Data Attributes.

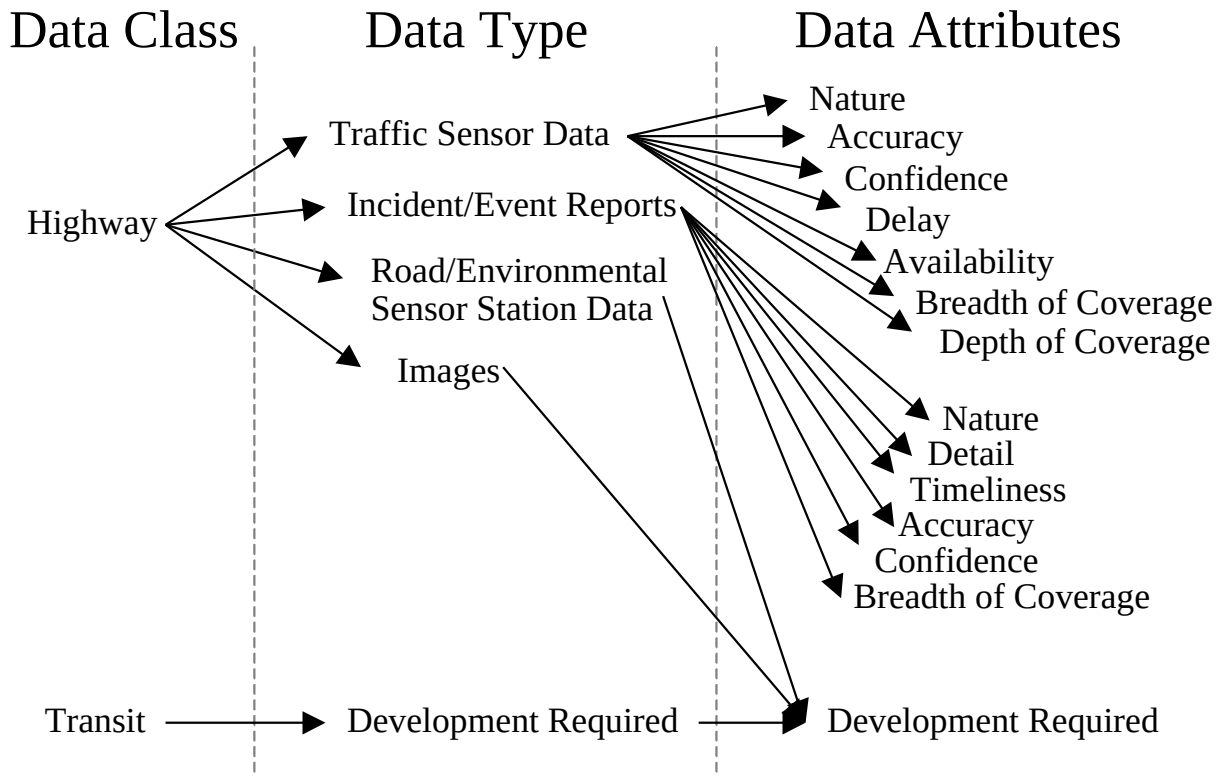
Data Types

Data Types are distinct classes of data necessary to collect to support traveler information services. In this document, we will discuss four data types identified presently as important to ATIS:

- Traffic Sensor Data
- Incident/Event Reports
- Images
- Environmental Sensor Station Data

Data Attributes

Data Attributes are definable and measurable parameters for a given data type. These attributes should address the scale, quality and utility of the data being collected. Attributes have been fully defined and documented in this report for Traffic Sensor Data and Incident/Event Reports. Attributes for Images and Environmental Sensor Data are offered for reader consideration, but do not have the same level of maturity and consensus support as the previous two data types. The figure below illustrates the relationship of data types and data attributes.



Traffic Sensor Data

Data of this type are speed, travel time, volume and occupancy data or other numerical measurements used to characterize the flow of vehicles at a specific point or over a specific segment of roadway. This data can be generated from many types of detection systems, such as loop detectors, microwave, infrared or sonic detectors, video image detection, automatic vehicle identification, license plate matching systems and wireless phone probes.

Attributes

Nature

What data parameter is being collected? Four parameters can be used when collecting traffic sensor data for traveler information:

- *Volume* – the actual number of vehicles observed to be passing a point during a given time interval
- *Occupancy* – ratio of the time that vehicles are present at a detection station in a traffic lane compared to the time of sampling
- *Speed* – a rate of motion, as distance per unit time
- *Travel Time* – elapsed time for a vehicle to traverse a roadway segment

Volume, Occupancy and Speed data are typically collected at a point in the roadway. We define this as point data. Travel Time data must be collected over a section of a roadway (as travel time is the time is takes to move from one point to another). We define this as section data. Section data can also be described as link data or segment data.

Accuracy

On average, how closely does the data collected match actual conditions? All traffic sensor data collection systems are subject to inaccuracies. For example, loop detectors in adjacent lanes could both count a vehicle in the process of changing lanes, resulting in double counting the vehicle. Other technologies are subject to weather conditions, radio frequency interference, and occlusion. Clearly, the more accurate the data, the higher quality. This attribute is typically characterized using percentages. Often error is used in place of accuracy. For instance a system may be 90% accurate, or have a 10% error. This attribute is used to describe the average performance of the traffic sensor, not the accuracy of a precise data element.

Confidence

Is the data trustworthy? This attribute is a microscopic analogue to the Accuracy attribute. This attribute describes the degree of specificity of the system's belief in the quality that is communicated along with each data record. The more a data collection system can flag data as possibly being anomalous, the more likely a user of the data can ferret out bad data before using it. An example of bad data is the not too uncommon situation where a loop detector is "stuck on", or reports occupancy at 100% and zero (0) volume for consecutive measurements. Unless a single vehicle is sitting over the loop during an entire data-sampling period, this data are false.

Delay

How quickly is the data collected available for use in ATIS applications? It takes some amount of time for data to be gathered into a useable form at the roadside. It also takes time to communicate a sensor's data to some location for the processing necessary to fuse data from multiple traffic sensors. Performance vs. cost tradeoffs are a significant factor in data collection system design. Clearly, the fresher the data, the more useful it is for ATIS. But quicker data may require greater investment. Delay is usually characterized in seconds or minutes. Timeliness and Latency are synonyms of Delay.

Availability

How much of the data designed to be collected is made available? While traffic sensor data collection systems are usually designed to operate continuously, inevitably a user of the data will lose access from time to time. This attribute describes the average probability that a given data element will be available for use from a given traffic sensor. For example, if a radar detector was reporting average speeds at a specific point over five minute intervals, 12 data points will be generated each hour, 288 data points each day and 105,120 data points each year (105,408 in leap years!). If 2,100 data points were not available for use over the course of a year, the availability would be 98%. This attribute essentially combines factors such as sensor reliability, maintenance responsiveness, and fault tolerance into a single measure related to data output. The better the traffic sensor data collection system is designed, operated and maintained, the higher the data availability.

Breadth of Coverage

Over what roadways or portions of roadway are data being collected? If a region or roadway has multiple attribute characteristics, it is necessary to define each separately, indicating the coverage that each applies to. In other words, if an old traffic sensor data collection system covers one roadway operated by an agency and another roadway operated by the same agency has a state-of-the-art traffic sensor data collection system, it is likely they will have quite different attribute characteristics describing their performance.

Depth of Coverage: Density

How close together/far apart are the traffic sensors? For point data collection sensors, this attribute is characterized by sensor spacing. The closer together point sensors are, the greater the chance to develop an accurate description of the traffic situation. For section data collection sensors, this attribute is characterized by section length. The shorter the section length, the easier it is to develop aggregate travel times to and from different sections.

Incident/Event Reports

This data type is characterized by descriptive information on planned or unplanned occurrences that affect or may affect traffic flow. Incidents, construction/maintenance, events, road conditions and weather conditions are some of the types of data collected. This data are usually manually entered into a "system," although it can be stored and communicated either a text or through numeric codes. The manual entry into a system is the key differentiation from the Traffic Sensor data type

Attributes

Nature

For what type of incident/events are data collected? Using the extensive standards work in this area as a guide, it is possible to gather data on several types of road-related incidents/events:

- Crashes, breakdowns or other unplanned vehicle stoppages
- Planned or emergency roadway construction or maintenance
- Special events
- General road conditions
- General weather conditions
- Traffic control device malfunctions
- Disasters

Detail

What kind of data are collected associated with an incident/event? While it is possible to gather an infinite array of data pertaining to an event, generally desired data will fall into the following categories:

- *Reason* – description of the cause of the incident/event
- *Location* – where the incident/event is
- *Severity* – the magnitude of the incident/event
- *Impact* – the effect the incident/event is having on the roadway
- *Status* – current state of the incident/event
- *Advice* – suggestions for vehicles/travelers impacted by the incident/event
- *Duration* – anticipated length of time an incident/event will impact the roadway
- *Time* – the time data are provided regarding an incident/event

The more data collected about an incident/event, the more useful it is. Often overlooked in incident/event data collection is documenting the duration of an incident/event and the time a report is inserted into a "system." It is not uncommon for data regarding an incident/event to be collected and shared, but data indicating the end of an incident/event is not provided. Also, a time-stamp on data aids those using the data to determine the currency of the data.

Timeliness

How long does it take to detect, verify and update incidents/events? There are three distinct stages of incident/event data collection of interest for ATIS purposes:

- *Detection*: something might be out of the ordinary, but the data collector cannot yet be sure if an incident/event is occurring or has a complete data set regarding an incident/event.
- *Verification*: As is it determined an incident/event is occurring, accurate and complete data are collected
- *Updates*: Incidents can last for a few minutes up through a few years. It is necessary to provide updated data as the incident/event evolves.

Accuracy

How accurate is the data associated with an incident/event? The accuracy of an incident/event report will be determined by the correctness of the details reported. For example, a report will be in error if the impact is not appropriately quantified.

Confidence

To what degree is the data from an incident/event verified? Incident/Event reports can be generated from multiple sources including the monitoring of roadside surveillance infrastructure, cellular call-in reports, and reports from another agency or ISP. Hence, the confidence of the report will depend largely upon the reliability of the information source.

Breadth of Coverage

Over what roadways or portions of roadways are incident/event data being collected? Regions or roadways with multiple attribute characteristics should be defined separately indicating the varied levels of coverage.

Additional Data Types

Additional data types for future consideration for guidelines include image and Road/Environmental Sensor Station data. Presently, consensus efforts on these topics are not mature enough for inclusion in ATIS-related data collection/sharing guidelines. The initial thoughts as understood to data are summarized below to continue to build consensus on these topics.

Images

This data type represents a snapshot of a roadway to give a visual depiction of current traffic conditions. Images give a quick impression of traffic conditions that can be easily assessed, which consumers of ATIS services find valuable. However, this data type is not conducive to deriving detailed information such as that can be provided by traffic sensors. Images can be disseminated through multiple outlets including web pages, TV stations, kiosks, etc. Images can also be used by ISPs to verify or identify information that can then be manually inserted into a traveler information service.

Attributes

Breadth of Coverage

Over what roadways or portions of roadways are images being collected?

Depth of Coverage: Density

How close together/far apart are the cameras? This attribute is characterized by camera spacing, which is a function of several parameters such as the alignment of the roadway, visual obstructions that could lead to occlusion (e.g. skyscraper), pole height, and the field of view/zoom of the camera lens.

Resolution

How grainy is the image? This attribute is measured by the number of horizontal and vertical lines broadcast in the image. The EIA (Electronics Industry Association) standard for television broadcast is 525 lines of resolution. Most CCTV systems will specify a minimum of 460 lines of resolution.

Refresh Rate

How often is the image updated? This attribute will be a function of communications bandwidth available between the source of the image and the consumer. Full motion video, which would be the optimum refresh rate, requires 60 frames per second. Many web pages currently displaying camera images are updated only once every minute.

Road/Environmental Sensor Station Data

This data type encompasses a wide array of sensors including those that monitor weather, roadway, surface, and air/water quality conditions. These sensors provide roadside data to support Road/Weather Information Systems (RWIS). Providing travelers with environmental data allows more informed selection of modes, routes, and departure times, resulting in improved safety and increased convenience for travelers. The utility of this data will vary with local environmental conditions.

Attributes

Nature

What types of roadway/environmental data are collected? Using the standards work in this area as a guide, it is possible to gather data on several types of roadway/environmental data:

- *Elevation/Atmospheric Pressure*
- *Wind Data*: direction, speed, gust direction, gust speed
- *Temperature*: air, wet-bulb, dew-point, 24-hr maximum, 24-hr minimum
- *Humidity/Precipitation*: relative humidity, adjacent water depth, adjacent snow depth, roadway water depth, roadway snow and packed snow depth, precipitation indicator and type, precipitation rate, snowfall accumulation rate, ice deposit (thickness), precipitation start time, precipitation end time, total precipitation past X hours

- *Radiation*: solar radiation, total sun, cloud cover situation
- *Visibility*: surface visibility (measured in tenths of a meter), visibility situation (clear, fog, smoke, sea spray, blowing sand/dust, sun glare, insect swarms)
- *Pavement Sensing*: pavement type, elevation, solar exposure, surface status (dry, moisture trace, wet, chemically wet, ice warning/watch, snow warning/watch, absorption, dew, frost), surface temperature, pavement temperature (2-10 cm below surface), surface water depth, surface salinity, surface conductivity, pavement freezing point, surface black ice signal, subsurface type, subsurface temperature, subsurface moisture
- *Pavement Treatment*: number of treatments, treatment type/mix (sand, dirt, gravel, cinders, water, salts, etc.), treatment form (dry, pre-wet, liquid), treatment amount (kilograms per lane kilometer), treatment width
- *Air Quality*: carbon monoxide, carbon dioxide, nitrous oxide, nitrous dioxide, sulfur dioxide, ozone, particulate matter
- *Water Quality*

Breadth of Coverage

Over what roadways or portions of roadways is roadway/environmental data being collected?

Depth of Coverage: Density

How close together/far apart are the sensors? This attribute is a function of local parameters.

DATA QUALITY LEVELS

A simple hierarchy of quality levels was developed to assess of the data attributes introduced for the traveler information data types. This hierarchy was developed from the notion of what is "good", "better", and "best" for each attribute. A "good" quality level suggests the minimum level of data collection that should be designed for each attribute. "Better" and "best" quality levels are indicative an improved level of service.

The "good" quality level was developed for each data attribute for each data type at a minimum. Good baseline data collection supports multiple traveler information services such as real-time monitoring, detection of incidents, and prediction of information. Hence, having "good" data corresponds to providing valuable services to consumers, thereby encouraging market growth and maturity. It should be noted that consumer expectations of data quality will change over time -- the more they see, the more they want. Providing "good" data provides a foundation service level for implementers and planners. As consumer expectations increase, "better" and "best" quality levels can be implemented to meet the demand.

Where appropriate, the distinctions are made between limited access highways and arterials. This addresses cases when a single quality level cannot address the different operational nature between the two types of facilities. Traffic sensor data collection on arterials for ATIS use less developed than for limited access highways. One assumption implicit in these guidelines is that point data on arterials is not – and will not be in the near future – capable of generating useful data for ATIS applications.

The section of this report, *Applying the Guidelines*, describes how an organization/region could use the quality levels to support their activities.

Traffic Sensor Data Quality Levels

Nature – Limited Access Highways

Good	*Aggregated Point data* – Data collected at a point. Data from individual sensors can be aggregated, but data from general purpose lanes should not be mixed with High Occupancy Vehicle (HOV) lane data
Better	*Discrete Point Data* – Data collected at a point. Data from individual lanes are provided without aggregation. ***or*** *Aggregated Section data* – Data collected over a section or segment of roadway. Data from discrete link measures can be aggregated, but data from general purpose lanes should not be mixed with HOV lane data
Best	*Discrete Section data* – Data collected over a section or segment of roadway. Discrete link measures, such as the travel times for all vehicles detected, are provided. Note that discrete data should not be associated with a vehicle or person to preserve privacy.

Nature – Principal Arterials

Good	*Aggregated Section data* – Data collected over a section or segment of roadway. Data from discrete link measures can be aggregated.,
Better	*Discrete Section Data* – Data collected over a section or segment of roadway. Discrete link measures, such as the travel times for all vehicles detected, are provided. Note that discrete data should not be associated with a vehicle or person to preserve privacy.
Best	*N/A*

Accuracy

Good	10-15% error
Better	5-10% error
Best	<5% error

Confidence

Good	*Qualitative Description* – Tiered Confidence Description - e.g., "good", "suspicious", "bad"
Better	*Quantitative Description* – % Confidence Factor - e.g., 95% confident
Best	*N/A*

Delay

Good	2-5 minutes
Better	1-2 minutes
Best	<1 minute

Availability

Good	90-95%
Better	95-99%
Best	> 99%

Breadth of Coverage

	Limited Access Highways	***Principal Arterials***
Good	Major Limited Access Roadways	Major Principal Arterials
Better	All Limited Access Roadways	All Principal Arterials
Best	N/A	

Depth of Coverage: Point Data

	Limited Access Highways	***Principal Arterials***
Good	Between Major Interchanges	N/A
Better	Between Every Interchange	N/A
Best	Maximum 0.5 mile spacing with at least one sensor site between every interchange	N/A

Depth of Coverage: Section Data

	Limited Access Highways	***Principal Arterials***
Good	Between Major Interchanges	Between Major Principal Arterials/Limited Access Highways
Better	Between Every Interchange	Between Every Principal Arterial/Limited Access Highway
Best	N/A	Between every signalized intersection/interchange

Incident/Events Data Quality Levels

Nature

Good	Crashes, breakdowns, or other unplanned vehicle stoppages Planned or emergency roadway construction or maintenance Disasters
Better	**Good +** Special Events General Road Conditions General Weather Conditions
Best	**Better +** Traffic Control Device Malfunctions

Detail

Good	Reason Location Severity Time
Better	**Good +** Impact Status
Best	**Better +** Duration Advice

Timeliness

Good	2-5 minutes
Better	1-2 minutes
Best	< 1 minute

Accuracy

Good	10-15% error
Better	5-10% error
Best	< 5% error

Confidence

Good	Verified, non-visual
Better	Verified, visual
Best	N/A

Breadth of Coverage

	Limited Access Highways	***Principal Arterials***
Good	All Limited Access Roadways	Major Principal Arterials
Better	N/A	All Principal Arterials
Best	N/A	All Principal Arterials + additional Arterials

DATA SHARING

One of the key issues to a data collection program is the sharing of data between the public and private sectors. Who has the data? Who needs the data? What kinds of policies are in place to maintain the integrity of the data amidst the data sharing? Is there a reluctance on the part of the public sector (which typically collects and maintains ATMS and APTS generated data) to share with the private sector? Is the data gathered by the public sector sufficient for private ATIS needs? Under what limitations will the public sector allow the private to use the regional data? What are the constraints in sharing data between public sector agencies or the private sector sharing their data with the public sector?

These and other questions will drive the guidelines and policies governing data sharing for individual public agencies.

The following sections outline issues on private access to public data for ATIS.

Issue Overview

A common complaint among private sector ISPs is that the public sector has data that they are not making available to them. Nearly as common are public sector complaints that such private sector requests are unreasonable, costly, or contrary to public policy considerations. Some states and metropolitan areas seem to have worked out reasonable agreements, but many others have apparently not done so.

Several approaches are possible, but the following steps are suggested as a starting point:

1. Gain a better understanding of the issues of all parties.
2. Identify the parties' specific objectives, how they vary, and what barriers exist.
3. Identify examples where those barriers have been creatively addressed.
4. Share the information in the form of recommended best practices, model agreements, etc.
5. Recommend other actions, if required

Issue Dimensions

There are many aspects of "data access," so we have to start to defining what is meant. To date, data and image seem to each have their own access-related issues. Some of the issues overlap and some are unique.

Access to Data

- Which data are to be made available (e.g., raw/processed, fine grain/aggregated, some/all, etc.)?
- How will the data be accessed (e.g., process and technology)?
- To whom will the data be made available (e.g., general public, registered ISPs/media, etc.)?

- How often will the data be accessed (e.g., continuous feed, periodic "subscription," etc.)?
- Is there a cost to provide access, and is there an appropriate cost recovery mechanism (if required)?
- Is there any system or privacy security issues? How will they be handled?
- Are there to be assurances (or caveats) to users regarding accuracy, reliability or availability?
- Should there be "performance" or data integrity requirements to ensure that the data are "properly" used?
- Is there any liability for use of the data, or relief from such liability?
- How would archived data be made available to interested public and private parties?
- Are there any other data issues?

Access to Video

- Which cameras will be available, how many simultaneously?
- Is there a choice of cameras, or are they predetermined?
- Can the private sector request pan/tilt/zoom control? If so, how is control to be managed?
- Is there a cost to provide access, and is there an appropriate cost recovery mechanism, if required?
- Are there any privacy or security issues? How will they be handled?
- Is there any liability for use of the data, or relief from such liability?
- Are there other video issues?

Public Sector Perspective

The private sector perspective, in its simplest form, is: "Give me all the data you have, raw or otherwise, real- or near-real-time, at no charge." The public sector issues seem to be more complicated. They include, but are not limited to, the following:

1. Equality of access: Public sector can't favor one private provider or lock out competitors (although some exclusive agreements have been negotiated).

2. Equality of service: Public sector wants at least a basic level of services available to all taxpayers/citizens, not just for high-income gadget-buyers.

3. Data integrity: Public sector doesn't want to give out data that are (1) of unknown or questionable quality; (2) may be easily misinterpreted or misused by naïve ISPs who don't understand the transportation data; or (3) likely to mislead or misinform the public.

4. Data selection: Public sector doesn't want to have to customize per each private user request (loop level, 15-second, lane by lane, cleaned, error-checked, processed, aggregated, etc.), prefers one-size-fits-all.

5. "Performance" requirement: Public sector wants to ensure that the traveling public is receiving the most up-to-date, accurate info, wants public safety and other critical messages to "get through" to end-users with little or no modification.

6. Costs: Public sector doesn't have staff or resources to respond to multiple custom requests; doesn't have systems capable of accommodating unlimited numbers of simultaneous users.

7. Security: Public sector doesn't want confidential or sensitive information released; doesn't want to reduce system security (increased hacking attacks) by opening up systems.

8. Procurement: Public sector can't easily accept "gift" systems or upgrades to accommodate private sector requests.

Unresolved Questions to Date

The sharing of data already being collected has emerged through the development of these guidelines as a significant issue. It was determined to be an important enough issue to be included in this document. However, several issues remain to be resolved:

- Is there actually a problem in private access to public data? How widespread is the problem?
- Which issues are the most important, or most troublesome?
- Which public sector concerns are the most prevalent, or most problematic?
- Where have these problems been effectively addressed, and how were they worked out?
- What more should ITS America and FHWA do to deal with these concerns?

In the context of using this document to support data collection strategy development, the message to draw from this section is that it is important to assess data already being collected to determine if this data are being leveraged to the maximum extent possible for traveler information applications.

Guidelines

To underscore the importance of data access, a quality level has been established in this area. The quality levels are:

Access

Good	Licensed or open access, via a documented interface
Better	Licensed or open access, via a standard interface
Best	N/A

LOCAL FACTORS

Issues will undoubtedly arise when applying these guidelines to a specific region. While judgement should be used when applying these guidelines, we can offer some examples of situations that might require local modifications. Examples include:

Funding Availability – Ideally, implementers would be able to design their data collection systems to provide services characterized by the best quality level. However, funding resources will dictate an implementer's ability to provide services. As such, implementers should consider the development of an implementation plan to "step" towards the desired quality level using available resources.

High proportion of congestion on a facility is non-recurring – Such a facility will tend to have a greater variance of travel conditions from trip-to-trip, suggesting that information would be more useful than on facilities where travel conditions are more predictable. Using the *1999 Annual Mobility Report* produced by TTI as a guide, we would suggest that facilities that have 2/3's or more of the congestion as non-recurring are candidates for consideration for higher quality levels than suggested purely based upon congestion level.

A facility is a plausible alternative to another more strategic/congested facility – An example of this situation is in Broward County, Florida, where I-95 and Florida's Turnpike parallel each other as north-south limited access facilities a few miles apart the entire length of the county. I-95 has far greater daily congestion than the Turnpike. But due to the value of the Turnpike as an alternate route, particularly in extreme congestion on I-95, higher performance levels of data collection may be justified for the Turnpike than if examined independent of I-95.

If a region has a congestion index abnormally high or low for its size – Regions of the same size can often have quiet different levels of congestion. This could be due to terrain, weather, the abundance or dearth of capacity, or other factors. A region that has above average congestion for its size may consider adopting higher performance levels for its data collection system. Conversely, a region that has below average congestion for its size may choose a lower performance level.

Facility Congestion Levels – Departures from free flow conditions will occur on those facilities that experience high degrees of recurrent congestion. Such facilities will be more dynamic in character, thus driving the need for higher quality data collection. As such, facilities with extreme congestion levels will may require a greater level of quality of data to support traveler information services for that facility.

The point is that these guidelines should be used as a starting point for data collection planning, and local factors should be carefully examined to determine if changes need to be made to suit the region.

RELATIONSHIP TO OTHER ITS AREAS

ATIS is not a standalone delivery system, and as such, the relationship between ATIS data and data available and/or needed from other ITS cannot be overlooked. For most regional ITS, data are generated from ATMS and APTS, and these systems often drive the data collection needs. While ATIS can make use of this data, it is often not comprehensive enough to meet consumer (ISP, agency or motorist) ATIS data needs – hence, the gap.

As a key element of an integrated ITS, ATIS data collection efforts must be coordinated with the existing data collection of other areas and systems (ATMS, APTS, etc.). It is important for those developing a strategy for data collection to consider the needs of the various systems at work in their respective areas. This will help to foster a more coordinated regional effort toward data collection as well as a more cohesive plan for use and dissemination of the data from the various systems.

ITS America and other national consortiums (FHWA, AASHTO, ITE, etc.) are embarking on several efforts to develop standards and guidelines for data in various areas of ITS. These range from standardizing messages, nomenclature and terminology, as well as technology and communication protocols (see *Other Resources* for more information on some of these efforts). There is currently no standard for data collection, whether ATIS, ATMS or other ITS, and as a result, the data are not consistent even among similar systems in different regions. These concurrent data guidelines efforts in various areas of ITS will strengthen the activities of the individual data guidelines development.

ATIS needs / ATMS data

ATMS typically represents the majority of data collected in a regional ITS, usually as part of freeway and (some) corridor management systems. With the ATMS data collection technologies already deployed, there is usually a finite amount of information available from these systems. There is not a "standard" for ATMS data collection; as a result, the amount, type, and depth of data collected from ATMS will vary from system to system.

Consider the following questions when developing ATIS data collection needs vs. ATMS data needs or ATMS data currently available:

- What ATMS data are currently being collected? (data type – speeds, etc.)
- What is the coverage of the ATMS data collection (freeways, arterials – what are the boundaries?)
- Who "owns" the data? (agency)
- Is the ATMS data relevant to ATIS data needs? Is it of value to the ATIS data goals?
- Is ATMS data available or accessible for ATIS needs?

ATIS needs / APTS data

APTS data collection systems typically consist of automatic vehicle location (AVL) technologies on transit fleet vehicles. This data provides real-time vehicle tracking information to dispatchers for use in schedule monitoring, as well as bus travel speeds. The benefit of transit AVL data is that it could provide real-time information about arterial speeds not covered in a freeway data collection system.

Consider the following questions when developing ATIS data collection needs vs. APTS data needs or APTS data currently available:

- What kind of APTS data are currently being collected?
- What is the coverage area? (geographical as well as arterial/freeway)
- Who "owns" the data?
- Is the APTS data relevant to ATIS data needs? Is it of value to the ATIS data goals?
- Is APTS data available or accessible for ATIS needs?

ATIS needs / Transit-related Traveler Information Data

The amount and depth of transit-related traveler information will vary from region to region, depending on the transit activity and usage in a particular region. Much of the transit-related traveler information data collection would be covered in the APTS data (i.e., real-time vehicle information related to schedule, arrival times, etc.). While the primary users of this kind of information will be transit users, an evaluation of the types and availability of transit-related traveler information could determine its application in an ATIS data collection strategy. Types of transit-related traveler information could include park-n-ride lot information and availability, transit schedules, routes, service, etc.

Consider the following questions when developing ATIS data collection needs vs. transit-related traveler information data needs or data currently available:

- What kind of transit-related traveler information is currently being collected?
- Is this data of relevance to travelers? To a regional ATIS?
- Who "owns" the data?
- Is this transit-related traveler information available or accessible for ATIS needs?
- What data is of relevance to travelers making a pre-trip or en route mode decisions?

ATIS needs / Planning Data

Data that is used for planning purposes is typically historical data (accidents/incidents, traffic volumes, etc.), and as such, would not have application to an ATIS relying on real-time information. Planning data needs can, however, help those developing an ATIS data collection strategy by identifying significant routes that would be ideal candidates for data collection, or when considering the regional boundaries for a regional ATIS. The National ITS Architecture was amended in 1999 to provide for an Archived Data User Service (ADUS) being established. A process has been started to develop standards for data archiving. A draft strategic plan for ADUS Standards has been prepared as of the time of the writing of these guidelines. In the interim, it might also be helpful for those developing data archiving systems to consider what types of ATIS data would be needed for historical purposes and would facilitate identifying ATIS data collection needs.

ATIS needs / Other Systems

An ATIS data collection strategy should consider, from the outset, all available data from various systems in a region. In addition to ATMS and APTS, data from the following types of systems could be applicable to an ATIS data collection strategy (if available and depending on the region):

- Road/Environmental sensor station systems;
- Weigh-in-motion;
- Parking management systems (downtown business areas, airports, universities, etc.);
- Toll collection systems.

In all of the above areas, there are likely to be multiple sources or jurisdictions for the data. This will require partnering among institutions or public/private partnerships to evaluate data needs, availability, applications, and usage. See *Data Sharing* for more information about issues related to jurisdictional data sharing.

APPLYING THE GUIDELINES

Once the decision has been made that an ATIS data collection strategy should be put into place, it will be important for those planning the strategy to have a clear roadmap to reach the desired results. If an ATIS data collection effort is implemented, it is important to keep the long-term ATIS goals and needs in mind, as well as available funding for such a venture. An eight-step process of applying ATIS data guidelines is outlined below. These steps are intended to serve as general "guidelines," and should be tailored to suit regional ATIS data-specific needs and issues. Factors that could impact ATIS data collection, as well as deployment and funding considerations, are included where applicable.

1. Determine/define the region (boundaries) and stakeholders.

This would include geographic boundaries, agencies and jurisdictions, as well as intended ATIS consumers.

2. Outline "vision" for ATIS data to determine data needs and data priorities.

Having a long-term vision for the ATIS goals and system will allow stakeholders to evaluate data needs and priorities, as well as develop an incremental plan for ATIS data collection deployment. This vision will serve as a regional roadmap for ATIS development, and aid in long-range planning for ATIS as well as other systems that could contribute to ATIS data needs. This vision should also articulate the methods for data sharing.

3. Determine what types of data are currently being collected from other ITS.

In this step, consider the questions posed in *Relationships to Other ITS Areas* to determine what kinds of data are being collected by other regional systems,

- What kind of data are being collected by ATMS, APTS, or other systems in the region?
- What are the coverage areas for the existing data
- Can this data be used for ATIS?
- What jurisdiction owns this data?
- What are the coverage areas?
- Does this data fit within the ATIS data needs/vision?

4. Select desired corridors for coverage (factors).

In this step, all principal arterials and freeway corridors should be considered for potential ATIS data collection and coverage. Stakeholder coordination is particularly important at this step, as major freeways and principal arterials are likely to traverse jurisdictional boundaries. There are no national standards or guidelines to determine priority corridors; several factors need to be considered in the corridor evaluation and prioritization, and will need to be developed on a region-by-region basis. Some regions have short- or long-range transportation plans that identify corridors of regional significance; other regions will use statistics such as traffic volumes, accidents, or projected growth to determine the priority ATIS coverage corridors. Factors will need to be established and agreed upon by the regional stakeholders before consensus can be reached on the specific ATIS collection corridors.

Potential short-list of factors:

- Traffic volumes/recurring congestion
- Accidents
- Importance to regional mobility
- Anticipated growth
- Proximity to other corridors/alternate routes
- Others

5. Apply Minimum ATIS Data Collection Guidelines to selected corridors.

Using the *ATIS Data Collection Guidelines* identified previously, establish ATIS data collection objectives for each (or all) of the selected corridors. By outlining the goals for the selected corridors, there will be regional consistency as to the type, detail, and depth of the ATIS data collected, and this consistency will improve the overall regional ATIS program. Users will perceive consistent "quality" of information as a benefit.

6. Consider local factors and traffic management needs.

This step outlines factors that should be considered when applying the guidelines for data collection strategy development. "Local factors" will vary from region to region, and will include, at minimum:

- Existing data collection activities
- Congestion issues (regional and corridor-specific)
- Importance of corridor to regional mobility
- Proximity of ATIS candidate corridors to viable alternate routes
- Planned/future transportation infrastructure improvements
- Funding
- Other issues

These and other issues will require the input of stakeholders across the region.

7. Consider improved data quality goals.

Once the minimum ATIS Data Collection Guidelines are met there is the opportunity to enhance and improve the available ATIS data as funding becomes available. This can again be done by applying the same principles described above and by increasing the type, detail, and depth of the ATIS data collected. Improved ATIS data will improve the overall feasibility of the transportation systems and expands the general ATIS market.

8. Estimate cost and synchronize data collection strategy with funding profile.

Costs for the ATIS data collection strategy will need to be estimated based on several factors. These include:

- Infrastructure requirements (field hardware)
- Software (data retrieval, interfaces, archiving, etc.)
- Operations and maintenance
- Potential private sector funding

By making a preliminary estimate of cost, agencies within a region can get a feel for the funding requirements of the desired system, and, if necessary, make adjustments to the level of ATIS data collection deployment, coverage area, etc. For example, many metropolitan areas have data collection strategies in place on major freeways, but arterial coverage is lacking. The cost to deploy the same (or higher) level of arterial data collection could be out of range for many agencies, or left solely up to municipal or county jurisdictions who have ownership of the arterial corridors. By examining desired system and performance levels against estimated costs and available funding (existing and future), practical decisions can be made regarding data needs, coverage, and deployment timeframes.

A preliminary look at costs to deploy and maintain ATIS data collection could result in:

- Identifying partnering opportunities (inter-governmental agreements as well as public-private partnerships)
- ATIS mainstreaming with other planned infrastructure projects
- End-user fee-based ATIS considerations
- Others

OTHER RESOURCES

There are several resources that can be referenced to assist with developing an ATIS data collection strategy. With the emerging data guidelines, standards, and communications standards and protocols, the resources to be used for ATIS data collection is a dynamic entity.

ITS America and other transportation and technology consortiums (ITE, FHWA, etc.) recognized the need for developing data standards, and have embarked on several efforts to provide guidelines and standards for data collection and usage. Uniform guidelines that cover several areas of ITS will have a positive influence on moving toward regionally integrated and compatible systems.

ATIS Data Collection Workshop Proceedings

ITS America's ATIS Subcommittee has assembled proceedings from its ATIS Data Collection Guidelines Workshop held February 9-10, 2000. White papers from this workshop also can be found on-line at www.nawgits.com/jpo/atiswkshop.html. Information presented in these papers includes research on ATIS consumers, data needs, data "gaps" (between ATIS needs and other ITS data collection strategies such as ATMS and APTS), current barriers to ATIS data collection, and future directions for data guidelines development.

ITS America's ATIS Subcommittee has also assembled proceedings from its ATIS Business Models Workshop held October 1998. White papers from this workshop also can be found on-line and a Business Models Action Guide was published and is available from ITS America.

Standards

There are several standards development activities underway by various consortiums and national transportation organizations, including ITS America, ITE, FHWA, AASHTO, NTCIP, etc. The following is a brief list of standards activities that are currently underway.

Advanced Traffic Management Systems Data Dictionary

ITE, working cooperatively with FHWA and AASHTO led a national effort to develop a standardized *Advanced Traffic Management Systems Data Dictionary* (TMDD). The TMDD identifies and defines the specific data elements which make up the messages used within an ATMS and exchanged with other ITS applications such as APTS, ATIS and CVO. TMDD development has initially focused on high priority, high payoff areas within ATMS where interim products of standardized sets of data elements can be developed which will assist in accelerating system design and deployment. The final draft version of the TMDD has been approved by the ITE ITS Council Standards Review Panel. More information is available on ITE's website at www.ite.org on the Standards page.

Message Sets for External Traffic Management Center Communications (MS/ETMC2)
This standards publication was developed by a cooperative effort led by FHWA, ITE and AASHTO. This document provides a standardized set of messages for application in traffic management systems and associated information interfaces with other ITS services and functions. The development of standardized messages for ATMS is an essential part of the suite of ITS standards for information exchange and is an important step toward implementing the National ITS Architecture. More information is available on ITE's website at www.ite.org on the Standards page.

National Transportation Communications for ITS Protocols (NTCIP)
The primary objective of the NTCIP is to provide a communications standard that ensures the interoperability and interchangeability of traffic control and ITS devices. The NTCIP is the first protocol for the transportation industry that provides a communications interface between disparate hardware and software products. The NTCIP effort not only maximizes the existing infrastructure, but it also allows for flexible expansion in the future, without reliance on specific equipment vendors or customized software. There are several NTCIP standards development efforts underway, including data collection and monitoring, dynamic message signs, closed-circuit television, ramp metering, etc. To date, only the Dynamic Message Signs protocol has been published. Current NTCIP activities and committees can be found on www.ntcip.org.

Transit Communications for ITS Profiles (TCIP)
TCIP is a standards development effort designed to provide the interface structures that will allow separate transit components and organizations to exchange data. Importantly, the standard development effort produced a comprehensive set of TCIP interface requirements that allows effective and efficient exchange of data used for ITS user services and transit operations, maintenance, customer information, planning and management functions. The standard provides for interfaces among transit applications, which will allow users to communicate data among transit departments, operating entities such as emergency response services, and regional traffic management centers. Current TCIP activities and committees can be found on www.tcip.org.

ATIS Message Sets and Data Dictionary
The Society of Automotive Engineers (SAE) is developing a number of standards relevant to traveler information. Of particular importance are J2353, *ATIS Data Dictionary* and J2354, *ATIS Message Sets*. J2353 defines the data elements for ATIS messages and is the repository of unambiguous definitions needed to convey information to travelers. It provides the concise definition of data elements, including instructions on how to encode them at the bit level. It also describes the implied meaning of various phrases and points to other related data concepts on an element-by-element basis. J2354 provides the messages that are exchanged among information providers, traffic management centers, and other ITS centers and defines message sets for ATIS for general use independent of medium of transmission or bandwidth. The message sets themselves are made up of the data elements defined in J2353. More information is available on SAE's website at www.sae.org.

APPENDIX A

ATIS Data Guidelines Steering Committee

- Stephen Bahler, Minnesota DOT
- Michael Berman, Metropolitan Transportation Commission
- Harriet Chen, Traffic Station
- Christopher Cluett, Battelle
- Pete Costello, ITS America
- John Cox, Southern California Economic Partnership
- Ron Johnson, InfoMove
- Jane Lappin, EG&G/Volpe Center
- Mac Lister, Federal Highway Administration
- Joel Markowitz, Metropolitan Transportation Commission, ATIS Committee Chair
- Joseph Peters, U.S. Department of Transportation, Benefits, Evaluation and Costs Committee Secretary
- Dan Powell, PBS&J
- Pierre Pretorius, Kimley-Horn and Associates, Steering Committee Chair
- Robert Rupert, Federal Highway Administration, ATIS Committee Secretary
- Rick Schuman, PBS&J
- Valerie Shuman, Navigation Technologies
- Lawrence Sweeney, Etak
- Dale Thompson, Maricopa County DOT
- John Tolle, Federal Highway Administration
- Steve Wollenberg, Fastline
- Carol Zimmerman, Battelle

ATIS Data Guidelines Oversight Group

- Stephen Bahler, Minnesota DOT
- Les Jacobson, PB Farradyne
- Ron Johnson, InfoMove
- Joel Markowitz, Metropolitan Transportation Commission, ATIS Committee Chair
- Mark Nuaimi, Iteris
- Pierre Pretorius, Kimley-Horn and Associates, Steering Committee Chair
- Ivy Renga, Daimler Chrysler
- George Saylor, Ohio DOT
- Rick Schuman, PBS&J
- Bob Winick, Motion Maps